AD-A062 840   YALE UNIV NEW HAVEN CONN DEPT OF COMPUTER SCIENCE   F/G 5/7

INTERESTINGNESS: CONTROLLING INFERENCES.(U)

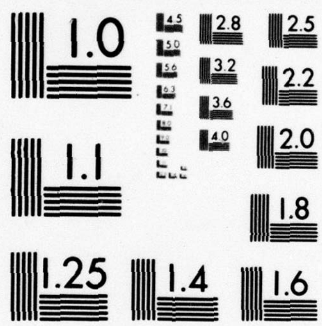
OCT 78 R C SCHANK   N00014-75-C-1111

UNCLASSIFIED   RR-145   NL

1 OF 1

AD
A062840

END
DATE
FILMED
2-79
DDC



MICROCOPY RESOLUTION TEST CHART

LEVEL II

12

INTERESTINGNESS: CONTROLLING INFERENCES

Research Report #145

by Roger C. Schank

October 1978

YALE UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

LEVEL II

(12)

INTERESTINGNESS: CONTROLLING INFERENCES

by

Roger C. Schank
Yale University
Department of Computer Science
New Haven, Conn. 06520

The research described here was done at the Yale Artificial Intelligence Project and is funded in part by the Advanced Research Projects Agency of the Department of Defense and monitored under the Office of Naval Research under contract N00014-75-C-1111.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

14 RR-

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 145
2. GOVT ACCESSION NO.: 9 Research rept.
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): 6 Interestingness: Controlling Inferences
5. TYPE OF REPORT & PERIOD COVERED: Technical
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): 10 Roger C. Schank
8. CONTRACT OR GRANT NUMBER(s): 15 N00014-75-C-1111
9. PERFORMING ORGANIZATION NAME AND ADDRESS:
   Yale University - Dept. of Computer Science
   10 Hillhouse Avenue
   New Haven, Connecticut 06520
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS:
    Advanced Research Projects Agency
    1400 Wilson Boulevard
    Arlington, Virginia 22209
12. REPORT DATE: 11 Oct[redacted]78
13. NUMBER OF PAGES: 43
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
    Office of Naval Research
    Information Systems Program
    Arlington, Virginia 22217
    12 49p.
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report):
    Distribution of this report is unlimited

    DISTRIBUTION STATEMENT A
    Approved for public release;
    Distribution Unlimited
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number):
    goals, inference, natural language, artificial intelligence, semantics, representation
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

The problem of controlling inferences is one of the most serious in AI. New types of goal and plan inferences seriously compound the problem. This paper attempts to outline one possible solution to controlling inferences, namely following what is interesting and ignoring what is not.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

407051

ABSTRACT

The problem of controlling inference is one of the most serious in Artificial Intelligence. New types of goal and plan inferences seriously compound the problem. This paper attempts to outline one possible solution to controlling inferences, namely following what is interesting and ignoring what is not.

B

-- OFFICIAL DISTIRUBTION LIST --

Defense Documentation Center Cameron Station Alexandria, Virginia 22314	12 copies
Office of Naval Research Information Systems Program Code 437 Arlington, Virginia 22217	2 copies
Office of Naval Research Code 102IP Arlington, Virginia 22217	6 copies
Advanced Research Projects Agency Cybernetics Technology Office 1400 Wilson Boulevard Arlington, Virginia 22209	3 copies
Office of Naval Research Branch Office - Boston 495 Summer Street Boston, Massachusetts 02210	1 copy
Office of Naval Research Branch Office - Chicago 536 South Clark Street Chicago, Illinois 60615	1 copy
Office of Naval Research Branch Office - Pasadena 1030 East Green Street Pasadena, California 91106	1 copy
Mr. Steven Wong Administrative Contracting Officer New York Area Office 715 Broadway - 5th Floor New York, New York 10003	1 copy
Naval Research Laboratory Technical Information Division Code 2627 Washington, D.C. 20375	6 copies
Dr. A.L. Slafkosky Scientific Advisor Commandant of the Marine Corps Code RD-1 Washington, D.C. 20380	1 copy

Office of Naval Research | 1 copy
Code 455
Arlington, Virginia 22217

Office of Naval Research | 1 copy
Code 458
Arlington, Virginia 22217

Naval Electronics Laboratory Center | 1 copy
Advanced Software Technology Division
Code 5200
San Diego, California 92152

Mr. E.H. Gleissner | 1 copy
Naval Ship Research and Development
Computation and Mathematics Department
Bethesda, Maryland 20084

Captain Grace M. Hopper | 1 copy
NAICOM/MIS Planning Board
Office of the Chief of Naval Operations
Washington, D.C. 20350

Mr. Kin B. Thompson | 1 copy
Technical Director
Information Systems Division
OP-91T
Office of the Chief of Naval Operations
Washington, D.C. 20350

Advanced Research Project Agency | 1 copy
Information Processing Techniques
1400 Wilson Boulevard
Arlington, Virginia 22209

Professor Omar Wing | 1 copy
Columbia University in the City of New York
Department of Electrical Engineering and
Computer Science
New York, New York 10027

Office of Naval Research | 1 copy
Assistant Chief for Technology
Code 200
Arlington, Virginia 22217

Captain Richard L. Martin, USN | 1 copy
Commanding Officer
USS Francis Marion (LPA-249)
FPO New York 09501

78 12 28 021

Interestingness: Controlling Inferences

1. Introduction

One of the central issues in natural language understanding for the last few years has been the problem of making inferences and controlling those inferences. Within our own group, the problem was brought to light most clearly by the work of Rieger (1973). Rieger's theory was that inference was a spontaneous mechanism. Thus, each of his sixteen classes of inferences was intended to fire off for every conceptualization (i.e. every state or event) that entered an understanding system. Since each new inference was also to be treated as an input to the system, it too would be submitted to the inference mechanism, thus creating a combinatorial explosion. Rieger's theory just didn't seem right.

One of the reasons it didn't seem right can be illustrated by an example I brought up which I later labelled as "the Muhammed Ali problem". Rieger's program, when presented with a sequence such as (1):

> (1) John was walking down the street when he spotted Mary. He walked over and hit her.

would begin to speculate on the state of Mary's health, and the reasons for John's actions. These inferences fired from the underlying conceptual action PROPEL.

The "Muhammed Ali problem" can be seen by substituting Muhammed Ali for Mary in the same story:

> (2) John was walking down the street when he spotted Muhammed Ali. He walked over and hit him.

Here, although Rieger's program would not detect the difference, most readers find themselves speculating about any number of possible inferences, the very last of which has to do with possible negative state changes of Ali's health. In fact, more likely in that regard, is to wonder exactly what might have happened next. That is, we are far more inclined to wonder about John's health. (If you do not get this idea from substituting Ali for Mary because of some knowledge about professional fighters, try substituting "a policeman" for Mary. The effect should be the same.)

Simply stated, the Muhammed Ali problem points out that certain inferences are more likely to be noticed than others in given contexts. That is, it seems clear that humans have some method of controlling inferences that allows them to focus on important things rather than unimportant things.

One of our first solutions to this problem was the idea of a script (Schank and Abelson (1977)). Scripts facilitate the processing of connected text (Cullingford (1977)), particularly text describing mundane stereotyped situations. We came up with the idea of scripts at least in part because of the problem of controlling inference set up for us by Rieger. It seemed obvious that when a hamburger was served in a restaurant, even though all

the inferences from PTRANS were certainly valid (the hamburger was now on the table; it was not where it was before; the waitress must have wanted the hamburger to be on the table; John must have wanted to control the hamburger, etc) they were clearly not the point. We found that we could control the inference process by using the information that main events in one scene of a script were usually strongly related to main events in subsequent scenes of the script. Thus, the most significant inference of "serving" is that it enables "eating". It would have taken Rieger's program a great many steps to infer that, but SAM (Cullingford's script applier mechanism) did it rather easily.

Sometimes however, the solution to a problem has as its maximum value the pointing out of new more important problems. This is what happened in this case. Scripts clearly account for only a very small part of our language understanding faculty. Few stories are entirely script-based. In investigating what underlies scripts, we came back to the original problem. We described in Schank and Abelson (1977) a theory of plans and methods to carry out those plans (planboxes and scripts). To explain plans, we began to investigate the goals that we might expect to give rise to them. At the same time as this theory was being developed, two of my students began writing computer programs to make use of those notions. The programs that they developed make heavy use of goals and plans.

An example of their input-output is shown below:

PAM (taken from Wilensky, 1978)

@RUN PAM

*(UNDERSTAND CD6)

THE STORY IS

JOHN WANTED BILL'S BICYCLE.
HE WENT OVER TO BILL AND ASKED HIM IF HE WOULD GIVE IT TO HIM.
BILL REFUSED.
JOHN TOLD BILL HE WOULD GIVE HIM FIVE DOLLARS FOR IT, BUT BILL DID NOT AGREE.
THEN JOHN TOLD BILL HE WOULD BREAK HIS ARM IF HE DIDN'T LET HIM HAVE IT.
BILL GAVE HIM THE BICYCLE.

The following is a question answering session with PAM:

*(DOQA)

QUESTION: Q1
Why did John walk over to Bill?
Because he wanted to get his bicycle.

QUESTION: Q2
Why did Bill give his bicycle to John?
Because he didn't want to get hurt.

QUESTION: Q3
What were the consequences of John's walking over to Bill?
This enabled him to ask him to give him Bill's bicycle.

QUESTION: Q4
What were the consequences of John's asking Bill to give him Bill's bicycle?
Bill told him that Bill wouldn't give him Bill's bicycle.

The following are paraphrases in English offered by Pam from the various actors' points of view:

*(TELL)

WHO SHOULD TELL THE STORY? *JOHN

I wanted to get Bill's bicycle. So I walked over to him, and I asked him to hand it over. He told me that he wouldn't hand it over. So I aksed him to sell it to me for five dollars. Then he told me that he wouldn't hand over his bicycle. I told him that if he didn't hand it over then I would break his arm. He handed over the bicycle.

WHO SHOULD TELL THE STORY? *BILL

John came over. He asked me to give him my bicycle. I wanted to keep the bicycle, so I told him that I wouldn't give it to him. Then he offered to buy it for five bucks. I wanted to keep the bicycle, so I told him that I wouldn't give it to him. He told me that if I didn't then he would break my arm. I didn't want to get hurt. So I gave him my bicycle.

POLITICS (taken from Carbonell, 1978)

****

Start of POLITICS interpretations and QA dialogs

****

INPUT STORY: +THE US CONGRESS IS EXPECTED TO APPROVE THE PANAMA CANAL TREATY.

*(INTERPRET US-CONSERVATIVE)

Q1: SHOULD THE US APPROVE THE TREATY?
A1: NO, THE TREATY IS BAD FOR THE UNITED STATES.

Q2: WHY IS THE TREATY BAD FOR THE US?
A1: THE UNITED STATES WOULD LOSE THE CANAL TO PANAMA
AND THE UNITED STATES WILL BE WEAKER.

Q3: WHAT MIGHT HAPPEN IF THE UNITED STATES LOSES THE CANAL?
A3: RUSSIA WILL TRY TO CONTROL THE CANAL.

Q4: WHY WOULD RUSSIA TRY TO TAKE OVER THE CANAL?
A4: RUSSIA WANTS TO EXPAND ITS MILITARY CONTROL.

THE PANAMA CANAL HAS HIGH MILITARY VALUE.

Q5: HOW MIGHT IT HAPPEN?
A5: THE UNITED STATES MAY NOT BE ABLE TO STOP ANYONE TAKING OVER THE PANAMA CANAL. TORRIJOS IS A COMMUNIST. TORRIJOS WILL LET RUSSIA CONTROL PANAMA.

Q6: WHY SHOULD THE US WORRY ABOUT RUSSIA TAKING OVER THE CANAL?
A6: BECAUSE RUSSIA WANTS TO CONTROL THE WORLD.

Q7: THEY, WHAT SHOULD THE US DO TO STOP THE SOVIETS FROM TAKING THE CANAL?
A7: CONGRESS SHOULD VOTE AGAINST THE PANAMA CANAL TREATY.

Q8: WHAT ELSE CAN THE US DO TO KEEP THE CANAL?
A8: THERE ARE NO OTHER GOOD ALTERNATIVES.

*(INTERPRET US-LIBERAL)

Q1: SHOULD THE US APPROVE THE TREATY?
A1: YES, THE PANAMA CANAL TREATY IS GOOD FOR THE UNITED STATES.

Q2: WHY IS THE TREATY GOOD?
A2: THE PANAMA CANAL TREATY IMPROVES RELATIONS WITH LATIN AMERICAN COUNTRIES. IT LETS EVERY COUNTRY USE THE CANAL.

Q3: WHAT HAPENS IF THE US KEEPS THE CANAL?
A3: THIS WILL CAUSE DISCONTENT AND CONFLICT IN PANAMA AND LATIN AMERICA.

Q4: WHY WOULD THIS LEAD TO A CONFLICT?
A4: THE CANAL ZONE IS A UNITED STATES COLONY. LATIN AMERICA DOES NOT LIKE COLONIALISM. THE UNITED STATES WOULD MAINTAIN COLONIALISM.

Q5: BUT, WHAT IF RUSSIA TRIES TO TAKE OVER THE CANAL?
A5: RUSSIA DOES NOT WANT CONFLICT WITH THE US. RUSSIA WILL STAY OUT OF THE CANAL IF THE UNITED STATES DEFENDS IT. THE UNTED STATES DEFENDING THE CANAL IS PART OF THE TREATY.

Q6: WHAT SHOULD THE US DO NEXT ABOUT THE CANAL?
A6: THE US SHOULD DO NOTHING AFTER CONGRESS APPROVES THE PANAMA CANAL TREATY.

While, I believe the work of Wilensky and Carbonell to be extremely impressive, there does exist an important flaw in the reasoning behind both of them. They both proceeded under the assumption that all goals and plans of all actors in a story must be constantly

monitored. It would not be appropriate to criticize that assumption for two reasons. First, the texts they designed their systems to work on were constructed such that this assumption was very useful; they were trying to investigate goal-based texts. The second reason is that the theory Abelson and I were designing encouraged exactly that; goals and plans had to be monitored all the time.

The problem with this assumption is that it leads you right back to the uncontrolled inference problem. If we must ask about the goals and plans of every actor we hear about, we will be concentrating on a great deal of useless detail.

Consider, for example, the following sentence.

(3) A small twin engine airplane carrying federal marshals and a convicted murderer who was being transported to Leavenworth crashed during an emergency landing at O'Hare Airport yesterday.

Taking seriously the idea of following all the goals and plans of the actors causes us to make goal and plan type inferences in (3) for the federal marshals, the murderer, the pilot of the plane, and possibly also the people in the control tower (to which "emergency landing" implicitly refers.) Thus we are left with inferring that the marshals have the goal of transferring the murderer to prison; that they have this goal because of a higher level goal involving a job; that this job is

a form of goal- subsumption (see Wilensky (1978)) for the goal of getting money; that they are acting as agents of the federal government; that the federal government has the goal of putting the murderer in jail and so on. We could also infer why the government wants to put the murderer in jail, why the pilot doesn't want his plane to crash, why the control tower doesn't want the plane to crash and go on in this fashion indefinitely.

We are left, in our new model of understanding, with a situation far worse than the one we were initially trying to correct. By inventing new classes of inferences to control the inference explosion problem, we have compounded our problem.

2. What Path to Follow

It seems unlikely that people actually think about why the federal marshals are doing what they are doing in (3). On the other hand, we cannot entirely ignore their goals either. If the story had been:

> (4) An airplane carrying two orangutans and a convicted murderer who was being transported to Leavenworth...

we would certainly not infer that the orangutans were taking the murderer to prison. That is, the SOCIAL CONTROL relationship between the marshals and the murderer is crucial and must be inferred.

But how much further in our inference path do we have to go? Need we infer why the marshals are in a SOCIAL CONTROL relationship? Should we begin to infer what their goals are? Both POLITICS and PAM say yes, but it seems obvious that at some point the answer must be no. We must control the inference process before it controls us.

The question before us then is to what extent we should follow a given path in a story. That is, which inferences are worth making and which are not? The answer is obvious - we need only make the inferences that help tie together the text. This answer was first proposed in Schank (1975) as a partial solution to the problem of controlling inference in a system for understanding stories. There are two problems with this answer. First, the ability to control inferences by tying together the text depends on our knowing before hand what inferences will lead naturally in that direction. How are we to know which goals of the federal marshals are irrelevant? We could spend time trying to tie together a text by processing all the goals in a story. So, this proposal solves nothing unless we know where to look for the information that will be of use in tying together text.

In Schank (1975) the needed information came from knowledge about causality. Tying together text meant constructing a causal chain of the states and events in the text. This could be done because we could easily store the normal enablements and results emanating from any given state or action. (Remember that we use a system of primitive states and actions and that there are very

few of these). Also, we had one event at a time to try and connect up to a preceeding event. But, on the higher level of goals and plans we do not know where to start. Whose goals and plans should we follow?

The second problem with this solution involves the actual mechanism for tying together events. While there may well be causal chains of events, it seems less likely that there are causal chains of goals and plans. What exactly does tying together mean here?

To return to our question: what inference path should we follow in understanding a story? Let me suggest a simple answer. We should follow the path that is interesting, of course.

This answer seems trivially obvious. Furthermore it depends on a definition of interestingness. What I shall attempt to do now is to define interestingness, motivate its usefulness in understanding, and write some rules for using interestingness as a method for controlling inferences.

3. Interestingness

What makes a story interesting? Consider the following:

(5) I went to Germany for my vacation. On one sunny day, I took a walk along the streets of Munich. First, I bought a lemonade from an old man. Then I saw some nice shop windows, an interesting statue, and visited a museum. I saw a nice hill and climbed up it only to

> find another hill which I also climbed. I was hot and tired and decided to find my way back. As I looked over the next hill before turning back, I saw in the valley below an army marching towards me. Behind me, I saw another army. Soon a battle broke out. I learned the next day that Turkey had invaded Germany. They had been repelled and the incident was hushed up. No one else knew about it.

In reading this story certain things are obviously more interesting than others. The "lemonade" and the "museum" are easily forgotten. On the other hand, the invasion is not easily forgotten. The last half of the story would be fascinating if anyone believed it. But why is it that the subject of believability doesn't come up for the first part of the story? Furthermore, why do there appear to be two distinct parts of the story?

The answer to all of these questions has to do with the concept of "interestingness". When something is deemed to be of interest, it is tracked by the reader. That is, we pay attention to it. What does pay attention mean? It means that we let our inference processes loose. Since this is nothing new, the import of this idea can only be that we are less willing to let our inference processes loose when a story or part of a story is uninteresting.

Where does this concept of interestingness come from and where will we find it during the course of processing? To further examine this consider the following sentences:

(6) John was walking down the street when suddenly:

a) he saw a cat
b) he decided to tie his shoe
c) he decided to eat a cookie
d) he decided to eat some chalk
e) he decided to drink the bottle of
Mr. Clean he just bought
f) he heard a loud noise
g) he heard an explosion
h) he heard an incredible boom
i) he heard a meow

To get a feel for what's going on here, it is worthwhile to attempt to characterize these sentences in terms of their potential interest. It seems obvious that a,b,c, and i are entirely uninteresting. In fact, they are a little weird, because the "suddenly" in (6) biases the reader to expect something interesting. On the other hand d and e are relatively interesting (e more than d) and similarly f,g, and h are increasingly interesting. This leaves us with a simple theory. First we have a relative rule: unusual things are more interesting than usual ones. And second, we have some absolute rules: noises are interesting, the louder the better, and the ill effects of eating inedible things are interesting, the deadlier the better.

These may seem like strange rules but they will do for starters. To consider it better, let's look at two stories that are structurally identical but vary in the interest value of the words (or concepts) contained within them:

(7) John was walking down the street eating an ice cream cone. He saw a man walk into the bushes and begin to undress. Soon a crowd had gathered and the police came to investigate. While they were there a giant explosion occurred two blocks away. People came running in their direction screaming that there had been a terrible accident. Many were bleeding and one man had lost an arm. Meanwhile a fire broke out in the

park. People said there was a conspiracy afoot because a bomb had been sighted nearby only yesterday. When an epidemic broke out the following week, everyone knew the aliens had landed.

(8) John was walking down the street eating an ice cream cone. He saw a man walk into the park and begin to read. Soon some pigeons had gathered and a boy came to feed them. While they were there a truck drove by a few blocks away. People who came walking towards the park said that it was a diesel truck. Many were hot and one man was really sweating. Meanwhile the park got really crowded. People said there was a new park being built nearby because a construction crew had been sighted only yesterday. When construction began the following week, everyone knew that the mayor had kept his promise.

Assuming that (7) is more interesting than (8), the question is why? To answer this, we must propose some absolutly interesting things. Consider one mentioned above, noise:

Noises are interesting because they can indicate impending catastrophe. Thus, when we read about an explosion in downtown we are interested, perhaps because we ourselves might have been there and might have been hurt. So, what really is interesting about noises is their potential harm. The first interesting thing we have then is DEATH.

1 - DEATH. Death is interesting (one would suppose) because everyone is worried about it happening to them. But, death is obviously not absolutely interesting. Few people consider the obituaries to be the most interesting part of the paper. This is because death, particularly at old ages, is usual and expected. This leads us to our second category of interestingness:

2 - UNEXPECTED EVENTS. The problem with using UNEXPECTED EVENTS as an interestingness category is that it implies a model of expectation which in turn means we will need a particular world model. Clearly what is expected for one person might not be for another. Now there is nothing wrong with that because what is interesting for one person might not be for another also. The problem is that while (1) was absolute (2) is relative to whomever we are talking about. What we can do here is to apply (2) as an operator on (1). Thus UNEXPECTED DEATHS are interesting.

This will work partially, but not completely. Obviously when we hear that John Smith died it is not expected if we did not know of his existence previously. It is also clearly uninteresting. There are exceptions to this however. If John Smith is, say 9 years old, and you have a 9 year old boy of your own, you might find it interesting. (Recall that interesting here means that we would want to know more about it, particularly by paying attention to the details and making appropriate inferences.)

John Smith's death might also be interesting if John Smith is a friend or relative of someone you know or of someone important, or if John died of some ailment that might possibly affect you.

We must add then to our absolute interest DEATH, and our operator UNEXPECTED EVENTS, a new absolute interest DANGER, and a new operator PERSONAL RELATEDNESS.

3 - DANGER. Danger is interesting because it is a possible prelude to DEATH. However to argue that this is why people find it interesting might miss the mark. People often put themselves in situations that simulate danger but where death is not a possible consequent. Sports fit well there. Also, there is the danger of losing what one values. This too is interesting and might have nothing to do with death.

4 - PERSONAL RELATEDNESS - Personal Relatedness is an operator, because, like UNEXPECTED EVENTS, it is really only the PERSONAL RELATEDNESS of another interesting thing that is interesting. Cutting your child's toenails is of strong PERSONAL RELATEDNESS but is not very interesting. Thus, PERSONAL RELATEDNESS alone is not necessarily interesting. But, as an operator on other concepts it can make things much more interesting than they inherently are. So, death of a friend or danger to a relative is interesting.

Other absolute interests are: POWER; SEX; MONEY (in large quantities); DESTRUCTION; CHAOS; ROMANCE; DISEASE and many other concepts and issues of this type.

There is also a kind of situational interest. This is the case where the items themselves are not necessarily interesting but their juxtaposition is. Thus, IRONY and DILEMMAS are interesting.

So what is interesting according to the above? If someone you knew died from having sex for a lot of money, that would be very interesting by my rules. Similarly if your house burned down that would be interesting, as would an epidemic in your neighborhood.

My point here is not to list out everything that could ever be interesting. That can and should be done through methods more rigorous than I have used here. My point is simply that people do have such metrics. The question now before us is: how do they use them?

4. The Use of Interest Values

In general, people are concerned with why they are being told a given piece of information. Thus, for every input an understander attempts to determine his need to know it. The major issue is how much work an understander should do in the way of inference at any given point. The problem then is really when to be top down and when to be bottom up. It is always neater to be top down, to have some overall framework into which we are trying to fit things and later to use that framework to interpret what we hear. But there is a problem with unexpected information or interesting information. Since understanders always notice such information, the question we must ask is 'how?' The detection of such information would seem to be a bottom up

process beceuse we cannot expect something to be unexpected or interesting. However, while assessing initial unexpectedness is certainly bottom up, the explanation for it is dependent on interest and relevance. Assessing interest therefore is a process which is both top down and bottom up.

Consider the following two stories:

(9)	(10)
Jonn went to visit Mary	Sam went to visit Fred
They had tea.	They had a beer.
He was about to suggest	Sam was about to suggest that
they go for a walk when	they play some basketball when
he spilled the tea on Mary.	he spilled the beer on himself.
He apologized and went home.	He went home.

How much plan, goal and other intentional inferences are appropriate in understanding these stories? In (9) we seem to recognize what John had in mind and feel for him. In (10) no such issues arise. Were we figuring out their plans and goals in either case? Another way of putting this is: When Sam visits Fred do we ask ourselves - Why? If not, then why do we in (9)? If we do, then how do we decide to stop looking for an answer? Why are we content to process on without answering?

One obvious answer here is that in understanding these stories, certain scripts are activated. Some of these scripts are simply more interesting than others. In particular, 'tea'

calls to mind the courting script and since the courting script has to do with ROMANCE it is to some extent interesting. This is not the case in (10). Nothing of interest comes to mind in the visiting a friend script.

The point then is that the inference of goals and plans may not be made unless we are interested in the characters intrinsically or in some of the circumstances that the characters find themselves in. The idea that we only make inferences when we are interested leads us to another issue that is obvious in the stating but somehow has been overlooked in AI research, namely the question of speed.

One of the reasons that we do not make all the inferences all the time may be that we simply do not have the time. Recall that input is coming in very quickly in spoken language and rather quickly in normal reading as well. One possibility is that we either figure it out fast or else we ignore it and continue on. The ramifications of this idea for how we do natural language processing are tremendous and it is worth digressing here for a bit to explore them.

5. Interest-Based Parsing

Simply stated the idea of interest-based parsing is this: It takes n microseconds to read a word and it takes m microseconds to completely process a word. (By completely process I mean, choose the correct meaning, integrate it fully into the overall representation of the text, and make the necessary inferences.) Since it seems quite likely that m is much larger than n in ordinary speech and reading, and since words come in continuous streams under those circumstances, then it is obvious that people cannot be fully processing every word they hear or read. What is more likely the case is that they are deciding what to pay serious attention to and what to pay casual attention to as they go. Such decisions can be explained on the basis of many factors, but the most obvious of them is interest.

Let us consider again sentence (3):

(3) A small twin engine airplane carrying federal marshals and a convicted murderer who was being transported to Leavenworth crashed during an emergency landing at O'Hare Airport yesterday.

It seems obvious that some parts of this sentence are more interesting than others. But more important than that, our ideas stated above concerning the processing time available would seem to imply that the processing of some words must take less time than the time it takes to read or hear them! Let's see why that is. We said above that it is likely that the time it takes to fully process some words must take more time than it takes to

hear or read them. Yet the total amount of processing time available is limited by the rate of the flow of input (which is continuous for speech). That is, it generally takes no longer to understand a sentence than it does to hear it. So, if we need to fully process some word because it is interesting or important, that processing must be done at the expense of processing other words at all.

The alert reader will have noticed that this argument assumes humans do not use any parallel processing capabilities in understanding. Even if they do, though, it seems to me that the total processing capability for use in understanding must have some bounds, and that the speed of input must often overwhelm those bounds. Thus we are still left with the necessity of processing some words at the expense of others.

The implication of this is that in many cases the understanding process is not entirely left to right. The understanding process must be top-down in order that the understander know what can be ignored and what is important. But when we process in a top-down fashion, the 'top' in a phrase may occur at the end of the phrase. The preceding words that had been ignored would then be 'gathered up', right to left. By right to left, I mean that while the stream of input is obviously left to right, some words are stored in a buffer and virtually ignored until a word that initiates processing is found. When such a word is found, the words in the buffer are gathered up and placed in their appropriate slots.

In order to process a noun phrase of the type that opens sentence (3) then, we must assume that a processor virtually ignores all the words until 'airplane', simply marking their existence in short term memory to be retrieved after the head noun is found. Once we know that 'airplane' is the subject of the sentence, expectations can be generated that allow us to have a better idea of what to look for (and therefore of what to ignore). Thus, 'carrying' can be virtually ignored because while we are only beginning to recognize what word it is, we have already heard about the marshals and the murderer and have decided to pay attention to those items.

The point here is that since we are really not seeing things one word at a time, but rather in a continous stream, we can pick out what we find interesting, go back to discover just those relationships that connect together what we are interested in and virtually ignore the rest. Do we care that the word 'carrying' was used instead of 'transporting', or that the construction used was not 'in which they were flying'? We have already predicted the relationship between the people and the airplane based on what it ordinarily is. We need only confirm the fact that nothing contradicts this prediction.

One of the consequences of this theory is that it should take longer, in normal reading mode, to read an interesting story than it does to read an otherwise syntactically equivalent dull one. Thus, (9) above should take longer than (10), (7) should take longer than (8) and sentence (3) should be faster than the

equivalent sentence in which the actors are riding on the plane rather than in it. (It might be rather difficult to actually get such results in an experiment however due to the problem that people might very likely slow down when bored and speed up when excited.)

A theory of interest-based parsing then, says that most words are barely noticed until some reason is found to pay attention to them. We must then ask what reasons there are to pay attention to a word and begin to process it seriously. Reasons for continuing to process a given word occur at all levels of the system. Some of these are:

Parser expectations: If the parser expects a certain kind of word, the satisfaction of that expectation can be taken as an extremely important force in the parser. Thus, a parser might function best that expected certain syntactic or conceptual types to the extent that it ignored everything else until it found them. This is again a violation of strict left to right parsing, but this violation makes sense considering our arguments above. (There is an aside here that is worth making at this point. We have talked over the years about how expectations drive various parts of the understanding process (Riesbeck, 1975 and Cullingford, 1977 for example). The contrast here is between expectation-driven processes and interest-driven processes. Obviously the most powerful and important mechanisms available to an understander are both expectation and interest driven at the same time.)

Syntax: Main nouns in a noun phrase can cause a processor to try to gather up all relevant modifiers. Also, certain function words can force a processor to pay attention if they might flag something important or expected.

Interest values: How does the parser decide what it wants to pursue? Obviously we need a fully integrated system where the parser and memory talk during the analyxsis of a sentence. Without such integration, the memory could not inform the parser of what to pay attention to. Interest values are stored in memory as part of the knowledge associated with concepts. As we have seen, certain concepts are absolutely interesting. Others are interesting only under certain circumstances (we will discuss these later on in this paper). Certain things are interesting on the basis of what has preceded them. Thus, the object of a shooting might be expected to be more interesting if the shooting took place in an embassy as opposed to a generally low-interest location such as a bar. (But of course, contexts can be created where bars are very important. This is why it is necessary to have memory available as opposed to just a dictionary.)

Top level expectations: If we are reading about an event that fits into a high level knowledge structure such as a script or a plan, predictions from that script or plan can focus interest during the processing of a sentence. Thus, we can know that the target of an assassination and the identity of the assassin are of critical importance in reading a story about an assassination and we can thus focus in on those items as top-down predictions

during parsing.

To see how all this is used we can return to our original examples (9) and (10): In a sentence "X went to visit Y", X is evaluated for interest value by memory because it is a head noun and because it is a person. When no information is found, the processing should be faster than when information is found. Thus, when X is 'John' or 'Sam' one proceeds quickly. If X were Henry Kissinger or one's mother one would presumably proceed more slowly because more expectations about their behavior would be found. (Since they are subjects of the sentence that is what is relevant here.)

'Went' is an item that urges us to continue processing since it has no meaning by itself. (That is, we could have 'went crazy', 'went fishing', 'went to Boston', and we can't do anything until we see the next few words. The point here is why speculate at all, just ignore it.) 'Visit' is a word that calls up a script ($VISIT) if the object of the visit is an equal or a family member. But remember that other scripts can be called up by the word visit that are distinct from $VISIT. If the object of visit is 'museum' or 'Congress' we would get quite different scripts or even no script at all. (What script does "went to visit a mortuary" bring up?) Again, I would claim that 'visit' is also almost totally ignored since it too means very little. Its real purpose is to get us very interested in the object that follows next. That is, we don't really start processing this sentence seriously until we see what Y is. Then, if Y is a

relative or friend then we instantiate $VISIT. If Y is a member of the opposite sex, we have an ambiguous sentence from a scriptal (and thus from a processing) point of view. In that case, either $VISIT or $ROMANCE would be applicable, and we will now have to figure out which.

We are thus ready, when the next sentence is seen, to try to find answers in story (9) that we don't care about in story (10). Furthermore, when "they had tea" is read and $ROMANCE becomes more likely, we must do still more processing. Recall that we originally started this discussion with an attempt at deciding the issue of when to track a character's goals. Clearly John's goals are more interesting than Sam's. This is because, as we have said, ROMANCE is interesting. Visiting obviously is much less interesing in a neutral context.

Tracking John's goals is done because they are easily discernable from the script he is in and the script he is in is marked as interesting. Thus when his script gets fouled up we feel for him in a way that we do not when Sam's script fouls up. The reason is simple: we were more interested in John.

Of course, much of what I have said here is quite speculative. But, if true, the ramifications for how we process language (on a computer and off) are tremendous. Since the crux of the matter is what is interesting and how those interest values are computed we will continue that discussion in the next section. The question of how interest values alter our concepts of parsing will be further discussed in Schank and Birnbaum

(1978).

6. Goal-Based Stories

PAM is a program that understands goal based stories. Wilensky (1978) cites many examples of stories his program can handle. These stories tend to have a form that includes a main character who has a goal, a problem in his attainment of that goal, (based on goal conflicts, or competition with other people), and the solution of that problem.

As long as stories are entirely goal oriented, Wilensky's program appears to be a major achievement in story understanding programs. But, the question of when goal based inferences are appropriate is a serious issue. Wilensky's theory shares many of the virtues and defects of Rieger's theory. Both present a theory and a program that demonstrate how an important class of inferences can facilitate understanding. But both somehow believe that their class of inferences ought to be made all the time. The question of the truth of that assertion still remains, although it is certainly more acceptable in Wilensky's case than in Rieger's since Wilensky's inferences are far less random and are structured in a good way. Still uncontrolled inferences can lead to very slow processors.

Hopefully, the ideas presented here with respect to interest should help to control the proliferation of inference problem. To see what I have in mind consider the following two PAM-like stories:

(11) John wanted to watch the football game. He started to turn on the TV when Mary announced she wanted to watch the ballet. Mary left the house crying and promised never to return. When she did come back John had bought her roses. They kissed and made up. It turned out the electricity had gone out everywhere. They laughed.

(12) John wanted to commit a mass murder with his shotgun. He started to load it when Mary announced that she wanted to go deer hunting. Mary left the house crying and promised never to return. When she did come back John had bought her a machine gun. They kissed and made up. It turned out that the invasion had started and there was no one around to kill. They laughed.

As far as I can tell, the processing for these two stories by PAM would be virtually identical. They are both stories that include a goal conflict based on limited resources and a resolution of that conflict. Yet, these stories seem different in an important way. The goals are of such a different character in (12) than they are in (11), that it is hard to believe that an understander would treat them identically. Mass murders are simply more interesting than watching television. Invasions are more interesting than power failures.

To see my point here, it is necessary to imagine these stories within the context of larger situations and explanations of those situations. The point here is not to criticize Wilensky. For contrived materials such as (11) and (12), the processing might be a great deal the same in either case. But, within a larger context, we would want to seriously track the consequences of the invasion and the mass murder. The goals and plans there are simply a lot more important to us than are the ones expressed in (11).

So, just as parsing considerations can be controlled by interest values such that not everything has strict attention paid to it, the same is the case with inferences. We should only make inferences to the degree that we find something interesting. When items of competing interest are around, we would choose the more interesting to infer (i.e. think) about. This brings to mind a point Wilks (1973) used to make about parsing. His preference semantics "preferred" one sense over another when it had the choice. No meaning was ever rejected, but in a choice situation one was preferred for reasons having to do with the context of the rest of the sentence. A similar phenomenon occurs here. In the absence of anything else to do, we infer on what we have. But, we prefer to infer about interesting things. So the same item (e.g. a power failure) that can get us thinking about how it affects the lives of those hit by it, will be ignored in the presence of an invasion. Thus, interests can be overriden by other interests. Overriding interests then, serve to override inferences, which serves to focus attention and speed processing, avoiding useless search.

7. Interests

In section 3 we sketched out some general categories of interest types. Here, we shall attempt to get more specific about what is likely to be interesting and what such an assessment of interestingness means.

a-Normality of States: The simplest and most obvious observation

to make is that something is interesting in direct proportion to its abnormality. Abnormality marks something as being interesting at all levels of understanding:

"John has a white house" is less interesting than "John has a red house with pink polka dots". Two questions are of interest here: 1- how do we know this, i.e. how is this information stored and retrieved? 2- what effect does this information have on actual processing?

1. The answer to the first question depends on our having a standard normality frame available to us during processing. Such a normality frame would have to work as follows: Normal behavior of people would be contained under frames (this is Minsky's (1975) sense of the term) organized around various aspects of their lives. Thus, there would be a NORMAL POSSESSIONS frame for example, that would be used as the basis for inferring the prototypical man's possessions. People are capable of answering questions about the prototypical man's posessions (e.g. does the average American own an elephant?, does the average American own a TV set?), so it seems obvious that such information must be available. Furthermore, it is absolutely necessary to have such information in order to understand a great many English words in their commonc senses. Thus, words such as 'ambitious', 'aggressive', 'hippie', 'obsequious' etc. cannot be understood except by comparison with a norm (this issue is discussed in greater detail in Schank and Lebowitz (1978) and in Carbonell (1978)).

Having this normality frame for possessions available allows us to look up the normative color of houses that people live in. Now by this I do not mean that there is only one normative color. Whatever is common would be listed there. Also, the existence of house on the possessions frame would make such a possession normative, causing us not to focus on it. The opposite would be the case for example in 'John has a purple helicopter' where neither the color nor the possession are normative. 2. When something has been determined to be non-normative ('abnormal' seems a bit loaded), it affects processing by causing us to focus on the non-normative (or interesting) item. Basically there are two kinds of such focussing of interests. When an item is deemed to be of interest, the actual practical effect of such a judgment is to cause the inference process to take a particular path. In general there are two kinds of paths available, forward and backward (see Schank and Rieger (1974)). Such paths are directions along a causal chain of events (see Schank (1973)). That is, either we will want to know the reasons for a given event or we will want to know the possible consequences of a given event. Such different interests will be marked for deviations from norms (as well as from other things). So, weird colors for houses are backward interesting (that is, why would someone do such a thing?) rather than forward interesting (not a lot follows from such a state of affairs, although some things do so it is to some degree forward interesting). The processing consequences then have to do with how many formward and backward inferences will be made. This depends on other things than just

the sentence itself of course. The surrounding context must be integrated into a connected causal chain in any story understanding task in any case (Schank (1975)), so the interest values serve primarily to give hints of where to look first in this case.

b-Normality of Actions: Actions have normality levels also. For example, if we hear that 'John entered the room by climbing in the window', this is more interesting than more usual entrances. 'Going down the chimney' is even more interesting still. To account for this we need to have normality descriptions at nearly every level. We need to know how people normally enter rooms, how they buy meat, how they get promoted and so on. This implies that any memory model must contain, in addition to its store of information about scripts, plans, goals and themes (Schank and Abelson (1977)), normality information with respect to each one of the specific knowledge structures that are known about. This in many cases is no more than the normal default values for a script for example. A script is constructed around such normality information and all we are saying here is that that information must be present for all types of knowledge structures and that that information is of great use in computing interest values. Its use is very simple. According to Rieger and Wilensky we need to be finding explanations all the time for the new information we receive. What I have been saying here amounts to this: we do not bother to find explanations for uninteresting events or states. Further, some things are interesting to the extent that we need to find such backward explanations, and some

are interesting to the extent that they cause us to look for expected consequences which may be of use in finding explanations for future inputs.

c-Concepts: Some concepts are interesting because of their relationship to the metrics proposed in section 3. By those rules concepts such as 'invade', 'bomb', 'explode', 'seduce', 'murder', and so on are interesting. To answer the two questions posed above, this information is stored in the dictionary under the concepts themselves and so it is in this case easily found. Its use is also found in the dictionary. So invade is both interesting forward and backward as are all the other words listed here. That means that we would follow stories involving such concepts more closely, trying to connect together other events we know about to the event that involves these concepts. I should point out that interest values are relative of course, so the presence of all these concepts in one story would cause those of lesser interest to be ignored and treated as if they were uninteresting. It should be remembered that what is at issue here is what to pay attention to and what to ignore, and that this can only be a relatively decidable issue. In a story describing the various methods of growing corn, a taller than normal cornstalk may be the most exciting thing around.

d-Scripts: Certain scripts are interesting and certain ones are not, just the way certain concepts are interesting and certain concepts are not. Thus, associated with a given script will be an interest value. For example, $CAR-STARTING,or $SECRETARY, are

not likely to be very interesting, whereas, $BOMBING, or $COURTING are more interesting. Here again these things can be listed in the memory under the script itself. The question of the effect of this interest value is somewhat more complex. When we say a script is interesting we can mean a number of different things by it. One thing we can mean is essentially the same thing that we have been discussing so far. That is, the consequences of that script taking place or the reasons for it taking place should be tracked and connected in a chain of events. Such an approach says that in one sense a script is viewed as an event in itself just like any other event. This view is correct but it is only one side of the question. Scripts are also structures inside which other events can fit. Because of this they can tell us things of relevance to the question of controlling inference by interests. In particular, they tell us whether to pay attention at all to events that fit inside the script itself. Sometimes we hear stories that relate all the events inside a script that we have deemed uninteresting (many of the initial stories that SAM processed would fit in here for example as would a great deal of everyday human conversation.) Since the script itself is uninteresting, it follows that an understander need not track with any great seriousness the events that comprise the script. That is, an understander in that situation need only decide that what he has heard belongs to the script. He need not worry about the potential inferences from an event once that has been decided. The only exception to this is if the new event does not match the normal flow of events in the

script or if the new event is itself deemed to be interesting on grounds found independant of its place in the script (there was poison in the hamburger for example). Although we did not realize it at the time, the question answering rules discovered by Lehnert (1977) that were shown to be often independent of the story itself can be used to argue that one need hardly pay attention to a normal script story at all in order to understand it. That is, if it is not interesting to read a story one need hardly pay attention to it and still not pay any penalty in question answering ability or understanding in general.

e-Plans, Goals etc: When we find that a script is of interest, then we know that it needs to be fit into a chain that connects it to other events. Such chains imply plans that connect to the script and goals that connect to the plans. Thus, one way that a plan can be determined to be interesting is if it has an interesting script or planbox (see Schank and Abelson (1977)) being used for carrying it out. A plan is rarely itself interesting, but it can be determined to be interesting by looking at what it is connected to. A D-CONTROL for example can be interesting if an interesting planbox (e.g. THREATEN or OVERPOWER) is being used; or if the object desired is of interest (e.g. the Kohinoor diamond); or if the goal dominating the plan is of interest (e.g. conquer Babylon, or become richest man in the world).

The consequences of determining that a plan is of interest are perhaps the greatest of all. When we determine that a plan is of interest, we can begin to intelligently follow what is going on in a story. We can track things that are related to that plan and ignore things that are not. Essentially then, all of this interest tracking is leading up to the time that we can safely look for items of interest in a story, knowing what to pay attention to and what to ignore. That is, we read looking for answers to 'did the murderer escape?', 'who fired the shot?', 'how did Nancy get pregnant?', 'What did the doctor say?'. Understanding then, is very much a process of tracking interests and knowing when to bring the full power of our inferential apparatus to bear upon the situation at hand, and, more importantly, knowing when not to.

f-Combinations of Concepts: It seems obvious that not only are individual concepts inherently interesting or not, but also combinations of uninteresting concepts can be themselves interesting. For example, 'eating' is not interesting, and 'Mr. Clean' is not interesting, but 'eating Mr. Clean' is. How to explain this is the topic of the next section.

8. Concept Combinations

The problem of combining concepts relates back to the Muhammed Ali problem. Neither hitting nor Ali are themsleves interesting but their combination is. Here, interestingness has a very important consequence. Recall that the original problem

was that there were a great many inferences that could be made but that it somehow did not seem that they all were being made. Did John's hand hurt? Was he angry at Ali? Did he have to move his hand? Any or all of the possible inferences could be correct, but somehow the question of Ali's response seems most significant. The question before us is 'why is this so'? The answer to this problem can be found by consulting a discrimination tree for the combination of concepts involved. What this discrimination tree does for us is tell us what is interesting and therefore what inferences should be followed. Let us consider the one for PROPEL since this is the action that underlies 'hit':

```
PROPEL (discriminations on the object of PROPEL)

first discrimination: dense or not dense
choose dense
second discrimination: capable of severely harming actor
                       of PROPEL or not
choose capable of harming
third discrimination: animate or inanimate
choose animate
           END OF BRANCH: Policeman; M. Ali; Murderer; Mafioso
           RESULTS: Interest value very high (10)
           INFERENCES: forward on harmful effects on actor
                       effected by object in new event tried first

choose inanimate container (includes vehicles)
           END OF BRANCH: car; building; train
           RESULTS: Interest value very high (9)
           INFERENCES:  forward on harmful state of actor resulting
                        from PROPEL tried first, harmful state of
                        people contained in object tried second

choose not capable of severely harming actor
fourth discrimination: depends on size of object and speed of PROPEL
           END OF BRANCH: wall; door; baseball
           RESULTS: Height of interest directly proportional
                    to speed and size combination
           INFERENCES: forward on harmful state of actor resulting
                       from PROPEL tried first
choose dense
fifth discrimination: personal effect or not
```

choose personal effect
sixth discrimination: friend or enemy
choose friend

    END OF BRANCH: list of positively related people
    RESULTS: Height of interest directly proportional
        to positive relatedness of person in object slot
    INFERENCES: backward on reasons for actions, forward on
        harmful emotional and relationship effects,
        then on physical effects.

choose enemy

    END OF BRANCH: list of negatively related people
    RESULTS: Height of interest directly proportional
        to negative relatedness of person in object
        slot
    INFERENCES: forward on physical effects and possible
        reprisals

choose no personal effect
seventh discrimination: potentially harmful to actor
choose not harmful

    END OF BRANCH: children, small animals etc.
    RESULTS: Height of interest dependent on value of object,
        children high, armadillos low
    INFERENCES: backward on reasons for action, forward on
        possible negative physical effects on object
        (but the lower the interest level the less
        this is done, thus, ants and pencils get
        little in the way of inferences.

choose possibly harmful

    END OF BRANCH: dogs, bears etc.
    RESULTS: Height of interest directly proportional
        to potential harmfulness of object
    INFERENCES: backward on previous events involving
        potential hazard to actor

The most important function of this tree then is to focus the inference process. Inferences are made because there is reason to believe that they will lead to interesting events. The decision to make an inference is precoded in this tree. Where this tree comes from originally is an open question, but something like its compiled form must be available to an inferencer.

To give a feel for what other trees of this type must look like, below is the INGEST tree given in a simpler form just listing a numerical interest rating. Any actual tree would have to look like the one above however.

INGEST

NOT POSSIBLE	NOT FOOD DEADLY	DEADLY OTHER FUNC	NOT FOOD NOT DEADLY FUN	NOT FOOD NOT FUN BAD TASTE	NOT FOOD NOT FUN NEUTRAL
air cond blackboard car INT= 10	arsenic INT=9	Mr. Clean Comet INT=8	coke dope liquor INT=7	chalk INT=6	dime marbles INT= 5

FOOD EXPENSIVE FANCY	FOOD EXOTIC	FOOD INEXPENSIVE GOOD	FOOD INEXP EH
caviar INT=4	poi taro root INT=3	steak lobster INT=2	lettuce hot dog INT=1

The presence of these discrimination trees implies that there is a set of values available for any given object (we ignore for the time being how these values become available). Thus, for an item such as chalk for example, we might find:

```
chalk  -  general INTEREST value (1)
          INGEST (6)
          PROPEL (2)
          GRASP  (1)
          PTRANS (0)
          ATRANS value (0)
```

These values mean that in general chalk will not be paid attention to. Some interest would be paid to chalk being eaten however. Zero values indicate that something will almost never

even be mentioned. If it is one might wonder why.

9.What to Pay Attention to

To see how to actually make use of what we have said so far, let us consider a few situations. Assuming John is someone we both know and I say "John is unscrupulous" it is reasonable to respond "Oh, what did he do?". It is reasonable to respond this way because unscrupulousness has a backward interest value of say 3 and the personal relatedness operator ups that to say 5. A backward interest value of 5 would be the condition necessary to cause a conversational rule of the type 'ask about previous event when a judgemental state has been expressed' to be fired off.

Now consider what would happen if I said "John is a murderer". We would expect an response of something like "What? What are you talking about? What did he do?". We are still backward interested here, only the interest is much higher. (We can express that by saying that murder is INT (10).) But murderer also has a extremely high forward interest value. This might be mitigated to some extent by the personal relatedness operator, but in general we are concerned that this murderer might threaten us as well. (To give a more common example of this, imagine being informed that your cleaning lady is a thief.)

So when something has a high forward and backward interest value, we in general want to know the details of past events and the prognosis for future ones. In story understanding, these are expressed as inferences up and down the causal chain. In

conversation they cause us to attempt to find out past details and worry about future events.

Now this is all not a simple process. It is mitigated in real life situations by a number of personal factors. For example if you have a positive affinity for John it is also necessary to infer forward about how he will escape the police and how you might help. In a story these inferences would be made as well but they would be secondary since the forward interest values for murder are lower than are the backward ones.

The methods I have described solve the Muhammed Ali problem:

> John was walking down the street. He saw Muhammed Ali.
> He went over and punched him in the nose.

Main inference: Was John hurt?

From the discrimination tree for PROPEL we get a forward interest value that is high that directs us to look at the effect of the event on the actor. This causes us to infer what Ali might do to John in retaliation which allows us to speculate on John's new physical state.

In addition to the combination effects that you get when two uninteresting concepts are present in the same event, we get combination effects when uninteresting scripts are interwoven. Thus, for example, undressing is not very interesting (say 2) and restaurants are not interesting (say 1), but, undressing in a restaurant is very interesting (say 6). Why? First, we know

that interference/resolution paths in a script are interesting (see Cullingford (1977)). Obviously there are various kinds of interferences. There are the standard ones involving lack of money or bad food or a restaurant that is out of food. There are also intrinsically interesting ones (that is, ones that are interesting without being embedded in another script) such as a fire or an earthquake. But there is also a class of combinatory interests. Undressing in restaurant is interesting because undressing is PRIVATE behavior that is being performed in a public place. This implies that all behaviour is marked for its public or private nature and that very private behavior in very public places is especially interesting. This we might mark privateness for the following: bathroom 10; sex = 9; sleeping 7, ...talking 0 . Here then, we would have such an interesting combination. Inappropriate behavior is interesting both forward and backwards, but it is mostly a forward process (then what happened?).

Public behaviour in a private place is also inappropriate but is less interesting. For example, asking for a menu at home or talking to your wife as you do to a student is inappropriate and mildly forward interesting. In addition, not conforming to role stereotypes in a situation is interesting as is deviating from any norm as we have stated before. (This of course brings us full circle since wealth, murder, ambition, aggressions are interesting and defined in terms of the norm in the first place.

At a higher level of knowledge structures certain goals are themselves interesting. For example, A (Achieve) - POSSESS is interesting while A - RESPECT is not. Similarly, A - LOVED ONES is not interesting and A - SUCCESS is.

Certain plans are interesting if they use planboxes that are at the high end of the scale (see Schank and Abelson (1977)). Thus OVERPOWER is maximally interesting while EXPRESS DESIRE is not. Weird ways of achieving a goal are also interesting as is not using a known planbox or violating a planbox's preconditions.

10. Conclusion

There is really very little to conclude at this point because this paper has been highly speculative. We can simply reiterate that it is crucial for us to understand how inferences are controlled. Perhaps interestingness and partial processing are part of the solution. In any event these ideas seem worth exploring both experimentally and computationally.

References

Carbonell, J.G. [1978]. Politics: Automated Idealogical Reasoning. COGNITIVE SCIENCE (2), pp. 27-51.

Cullingford, R.E. [1977]. Script Application: Computer Understanding of Newspaper Stories. Ph.D. Thesis, Yale University.

Lehnert, W.G. [1977]. The Process of Question Answering. Research Report #88. Yale University, Department of Computer Science.

Minsky, M. [1975]. A framework for representing knowledge. In P.H. Winston, ed. The Psychology of Computer Vision. McGraw-Hill, New York.

Riesbeck, C.K. [1975]. Conceptual Analysis. In R.C. Schank, ed. Conceptual Information Processing. North Holland Publishing Company, Amsterdam.

Schank, R.C. [1975]. Conceptual Information Processing. North Holland Publishing Company, Amsterdam.

Schank, R.C. [1973]. Causality and Reasoning. Technical Report #1, Istituto per gli Studi Semantici and Cognitivi, Castagnola, Switzerland.

Schank, R.C. and Abelson, R.P. [1977]. Scripts, Plans, Goals and Understanding: An Inquiry Into Human Knowledge Structures. Lawrence Erlabaum Associates, Hillsdale, New Jersey.

Schank, R.C. and Birnbaum, L.A. [1978]. Real-Time Integrated Parsing. Research Report #143, Yale University, Department of Computer Science, [forthcoming].

Schank, R.C. and Lebowitz, M. [1978]. Big Words. Research Report #144, Yale University, Department of Computer Science, [forthcoming].

Schank, R.C. and Rieger, C.J. [1974]. Inference and the Computer Understanding of Natural Language. ARTIFICIAL INTELLIGENCE (5), pp. 373-412.

Wilensky, R. [1978]. Understanding Goal-Based Stories. Ph.D. Thesis, Yale University, Department of Computer Science.

Wilks, Y. [1973]. An artificial intelligence approach to machine translation. In R.C. Schank and K. Colby (eds.) Computer Models of Thought and Language. W.H. Freeman, San Francisco.